CSC446 Homework #1, Kefu Zhu

1. Bishop 1.3

Suppose that we have three coloured boxes r (red), b (blue), and g (green). Box r contains 3 apples, 4 oranges, and 3 limes, box b contains 1 apple, 1 orange, and 0 limes, and box g contains 3 apples, 3 oranges, and 4 limes. If a box is chosen at random with probabilities p(r) = 0.2, p(b) = 0.2, p(g) = 0.6, and a piece of fruit is removed from the box (with equal probability of selecting any of the items in the box).

(A) What is the probability of electing an apple?

$P(Apple) = P(Apple|red) \cdot P(red) + P(Apple|blue) \cdot P(blue) + P(Apple|green) \cdot P(green)$

$= \frac{3}{10} \cdot 0.2 + \frac{1}{2} \cdot 0.2 + \frac{3}{10} \cdot 0.6$

$= 0.6 \cdot 0.2 + 0.5 \cdot 0.2 + 0.6 \cdot 0.6$

$= 0.58$

(B) If we observe that the selected fruit is in fact an orange, what is the probability that it came from the green box?

$P(Orange) = P(Orange|red) \cdot P(red) + P(Orange|blue) \cdot P(blue) + P(Orange|green) \cdot P(green)$

$= \frac{4}{10} \cdot 0.2 + \frac{1}{2} \cdot 0.2 + \frac{3}{10} \cdot 0.6$

$= 0.4 \cdot 0.2 + 0.5 \cdot 0.2 + 0.6 \cdot 0.6$

$= 0.54$

$P(green|Orange) = \frac{P(Orange|green) \cdot P(green)}{P(Orange)} = \frac{0.3 \cdot 0.6}{0.54} = \frac{1}{3}$

2. Bishop 1.11

By setting the derivatives of the log likelihood function (1.54) with respect to μ and σ^2 equal to zero, verify the results (1.55) and (1.56).

(1.54)

$$ln(p(x|\mu, \sigma^2)) = -\frac{1}{2\sigma^2} \sum_{n=1}^{N} (x_n - \mu)^2 - \frac{N}{2} ln\sigma^2 - \frac{N}{2} ln(2\pi)$$

(1.55)

$$\mu_{ML} = \frac{1}{N} \sum_{n=1}^{N} x_n$$

(1.56)

$$\sigma_{ML}^2 = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu_{ML})^2$$

Proof for 1.55

Calculate the derivative of the log likelihood function (1.54) with respect to μ

$$\frac{\partial}{\partial \mu} ln(p(x|\mu, \sigma^2)) = -\frac{1}{2\sigma} \cdot (-2) \cdot \sum_{n=1}^{N} (x_n - \mu) + 0 + 0 = \frac{1}{\sigma^2} \sum_{n=1}^{N} (x_n - \mu)$$

Maximize the log likelihood function by setting its derivatice to zero

$$\frac{\partial}{\partial \mu} ln(p(x|\mu, \sigma^2)) = \frac{1}{\sigma^2} \sum_{n=1}^{N} (x_n - \mu) = 0$$

$$\Rightarrow \sum_{n=1}^{N} x_n = N \cdot \mu$$

$$\Rightarrow \mu = \frac{1}{N} \sum_{n=1}^{N} x_n$$

Proof for 1.56

Calculate the derivative of the log likelihood function (1.54) with respect to σ

$$\frac{\partial}{\partial \sigma} ln(p(x|\mu, \sigma^2)) = -\frac{1}{2} \cdot (-2) \cdot \sigma^{-3} \sum_{n=1}^{N} (x_n - \mu)^2 - \frac{N}{\sigma} + 0 = \frac{\sum_{n=1}^{N} (x_n - \mu)^2}{\sigma^3} - \frac{N}{\sigma}$$

Maximize the log likelihood function by setting its derivatice to zero

$$\frac{\partial}{\partial \sigma} ln(p(x|\mu, \sigma^2)) = \frac{\sum_{n=1}^{N} (x_n - \mu)^2}{\sigma^3} - \frac{N}{\sigma} = 0$$

$$\Rightarrow \sigma \sum_{n=1}^{N} (x_n - \mu)^2 = N \cdot \sigma^3$$

$$\Rightarrow \sigma^2 = \frac{1}{N} \sum_{n=1}^{N} (x_n - \mu)^2$$

3.

Let (X ⊥⊥ Y) denote that X and Y are independent, and let (X ⊥⊥ Y | Z) denote that X and Y are

independent conditioned on Z (Bishop p. 373). Are the following properties true? Prove or disprove.

(A) (X ⫫ W I Z,Y) ∧ (X ⫫ Y I Z) ⇒ (X ⫫ Y,W I Z)

We have $P(X, Y, W \mid Z) = P(X \mid Y, W, Z) \cdot P(Y, W \mid Z)$

$\because X \perp\!\!\!\perp W \mid Z, Y$

$\therefore P(X, Y, W \mid Z) = P(X \mid Y, Z) \cdot P(Y, W \mid Z)$

$\because X \perp\!\!\!\perp Y \mid Z$

$\therefore P(X, Y, W \mid Z) = P(X \mid Z) \cdot P(Y, W \mid Z)$

Hence, we proved $X \perp\!\!\!\perp Y, W \mid Z$

(B) (X ⫫ Y I Z) ∧ (X ⫫ Y I W) ⇒ (X ⫫ Y I Z,W)

We have $P(X, Y \mid Z, W) = P(X \mid Y, Z, W) \cdot P(Y \mid Z, W)$

$\because X \perp\!\!\!\perp Y \mid Z \text{ and } X \perp\!\!\!\perp Y \mid W$

$\therefore X \perp\!\!\!\perp Y \mid Z, W$

$\therefore P(X, Y \mid Z, W) = P(X \mid Z, W) \cdot P(Y \mid Z, W)$

Hence, we proved $X \perp\!\!\!\perp Y \mid Z, W$